① الف) $P(r) = P.q + (1-P)(1-q) = Pq + 1 - P - q + Pq = (2P-1)q + (1-P)$

ب) $r = (2P-1)q + (1-P) \Rightarrow q = \frac{r+P-1}{2P-1}$

ج) $R = \frac{1}{n}\left[\underbrace{X(q_1)}_{\text{yes to } q_1} + \underbrace{X(q_2)}_{\text{yes to } q_2}\right] \Rightarrow E[R] = \frac{1}{n}\left[nPq + n(1-P)(1-q)\right] = Pq + (1-P)(1-q) = (2P-1)q + (1-P)$

$= \boxed{r}$

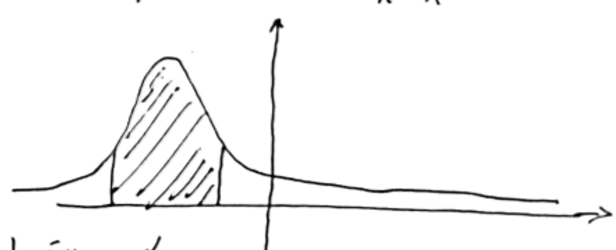
$\hat{Q} = \frac{R+P-1}{2P-1} \Rightarrow E[\hat{Q}] = E\left[\frac{R+P-1}{2P-1}\right] = \frac{E[R]+P-1}{2P-1} = \frac{r+P-1}{2P-1} = q$

د) $X \sim bernoulli(r) \Rightarrow R = \frac{1}{n}X \Rightarrow var(R) = \frac{\sum var(X)}{n^2} = \frac{n(1-r)r}{n^2} = \frac{r(1-r)}{n}$

ه) $Var(Q) = var\left(\frac{R+P-1}{2P-1}\right) = \frac{1}{(2P-1)^2} Var(R+P-1) = \frac{Var(R)}{(2P-1)^2} = \frac{1}{(2P-1)^2}\frac{r(1-r)}{n}$

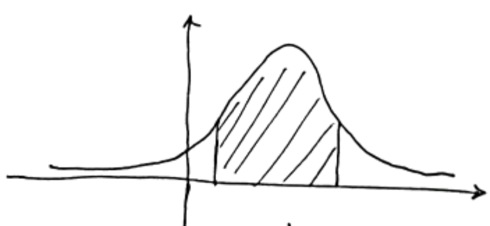
⑤ $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 \neq \mu_2 \end{cases}$

حالت $\bar{X}_1 - \bar{X}_2 < -t^*\sqrt{\frac{s_1^2}{n}+\frac{s_2^2}{n}}$



محدوده قبول H_0

حالت $\bar{X}_1 - \bar{X}_2 > t^*\sqrt{\frac{s_1^2}{n}+\frac{s_2^2}{n}}$



محدوده قبول H_0

در هر دو حالت میتوان $\mu_1 - \mu_2 = 0$ در ناحیه قبول H_0 نیست که به معنای reject شدن آن است.

۲) الف) $SE = \sqrt{\frac{pq}{n}}$ ، $SE < \frac{1}{100}$

$$p = \frac{3}{100} \Rightarrow n = \frac{(\frac{3}{100})(\frac{97}{100})}{SE^2} = 291 \quad , \quad p = \frac{4}{10} \Rightarrow n = \frac{(\frac{6}{10})(\frac{4}{10})}{SE^2} = 2400$$

باید اندازه نمونه ۲۴۰۰ باشد تا برای هر دو حالت $SE < \frac{1}{100}$ باشد.

$$\begin{cases} p = \frac{3}{100} \\ n = 2400 \end{cases} \Rightarrow SE = \sqrt{\frac{pq}{n}} = 0.0034 \quad , \quad \begin{cases} p = \frac{4}{10} \\ n = 2400 \end{cases} \Rightarrow SE = \frac{1}{100}$$

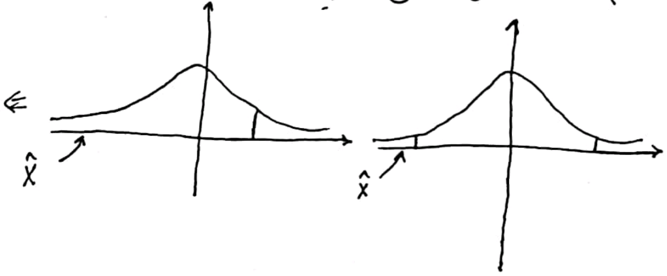
ب) $SE_1 < \frac{p_1}{100}$ ، $SE_2 < \frac{p_2}{100}$

$$\underset{p_1}{\Rightarrow} n = \frac{(0.03)(0.97)}{((0.01)(0.03))^2} = 3234 \quad , \quad \underset{p_2}{\Rightarrow} n = \frac{(0.4)(0.6)}{((0.01)(0.4))^2} = 150$$

۳)

الف) خیر، قضیه حد مرکزی همگرایی در توزیع می‌پردازد ولی بازه اطمینان می‌گوید با چه احتمال آماره مدنظر در بازه مشخص شده قرار دارد.

ب) درست، مثال نقض در پایین آمده است:



⇐ صرفا در سمت راستی reject می‌شود

⇐ به دلیل مشابه همواره ممکن است که ما حالت reject شدن را در نظر نگیریم.

ج) غلط، به طور معمول ما حداقل ۳۰ نمونه نیاز داریم تا بگوییم که sampling distribution را می‌توان با توزیع نرمال حدس زد.

د) درست، مقادیر outlier در توزیع های positively skewed باعث می‌شوند که مقدار mean از دو مقدار دیگر بزرگتر شود.

ذ) غلط، هرچه بخواهیم اطمینان بیشتری داشته باشیم باید بازه بزرگتری را پوشش دهیم. به همین سبب عریض‌تر خواهد شد.

ر) درست، دو عبارت داده شده هم‌ارز هستند و ما صرفا با محاسبه احتمال ۹۵ درصد بازه را مشخص می‌کنیم.

ز) غلط، بازه اطمینان ۹۵ درصد بیان می‌کند که ما در چه بازه‌ای به احتمال ۹۵ درصد آماره مدنظر را شامل می‌شویم و فارغ از تعداد نمونه‌ها است.

و) غلط، عبارت گفته شده درست است ولی مقدار اطمینان ۹۵ درصد گزارش نشده و ممکن است مقدار دیگری باشد.

ی) درست، عبارت گفته شده یکی از تعاریف بازه اطمینان است.

۴

$\bar{X} \to$ control میانگین گروه ~~...~~ $\bar{Y} \to$ میانگین گروه مورد درمان

$|\bar{X} - \bar{Y}| < 2 \to$ خواسته سوال $\Rightarrow \sim N\left(0, \sqrt{\left(\frac{100}{n}\right) + \left(\frac{100}{n}\right)}\right) \Rightarrow \int_{-2}^{2} N\left(0, \frac{10\sqrt{2}}{\sqrt{n}}\right) = 95\% \leftarrow$ s.t ، n

$Z_{\alpha/2} \cdot SE = 2 \Rightarrow 1.96 \cdot \frac{10\sqrt{2}}{\sqrt{n}} = \frac{4}{2} \Rightarrow n = 3.8416 \cdot 200 = \overset{768.32}{\cancel{\ldots}} \simeq \boxed{769}$

۶

$E[\bar{X}_c] = E[X_i] = \mu$, $E[\bar{X}_c] = E\left[\sum_{i=1}^{n} c_i X_i\right] \xRightarrow{E[X_i]=\mu} = E[\mu \sum c_i] = \mu \Rightarrow \mu E[\sum c_i] = \mu$

$\Rightarrow E[\sum c_i] = 1$

$\Rightarrow \sum_{i=1}^{n} c_i = 1 \Rightarrow$ $\bar{X}_c$ یک

تخمین زن unbiased خواهد بود.

۵

الف)

$\mu = np$, $\sigma^2 = npq \Rightarrow \begin{cases} \mu = \frac{50}{3} \\ \sigma = \frac{5\sqrt{5}}{3} \end{cases}$, $P(15 \leqslant X \leqslant 20) = P\left(\frac{15 - \frac{50}{3}}{\frac{5\sqrt{5}}{3}} \leqslant \frac{X-\mu}{\sigma} \leqslant \frac{20 - \frac{50}{3}}{\frac{5\sqrt{5}}{3}}\right)$

$= Z\left(\frac{2\sqrt{5}}{5}\right) - Z\left(-\frac{\sqrt{5}}{5}\right) = Z\left(\frac{2\sqrt{5}}{5}\right) + Z\left(\frac{\sqrt{5}}{5}\right) - 1 = 4871$

ب)

$E[X_i] = \frac{1}{6} \sum x_i = 3.5$, $Var(X) = E[X^2] - E[X]^2 = 15.16 - 12.25 = 2.91$

$P(X \leqslant 300) = P\left(\frac{X-\mu}{\sigma} < \frac{300 - 350}{10\sqrt{2.91}}\right) = P(Z < -2.94) \Rightarrow \varphi(-2.94) = 0.0017$

$\boxed{\text{Bonus Problem}}$

$$E[Z] = \int_0^T \frac{x}{\alpha} e^{-\frac{x}{\alpha}}\,dx + \left(\frac{1}{\alpha}\int_T^\infty e^{-\frac{x}{\alpha}}\,dx\right)T = \left[-\alpha x e^{-\frac{x}{\alpha}}\right]_0^T + \int_0^T \alpha e^{-\frac{x}{\alpha}}\,dx + T e^{-\frac{T}{\alpha}}$$

$$\Rightarrow \boxed{E[Z] = \alpha\left(1 - e^{-\frac{T}{\alpha}}\right) + T e^{-\frac{T}{\alpha}}}$$

$$E[Z^2] = \int_0^T \frac{x^2}{\alpha} e^{-\frac{x}{\alpha}}\,dx + \left(\frac{1}{\alpha}\int_T^\infty e^{-\frac{x}{\alpha}}\,dx\right)T^2 = 2\alpha^2\left(1 - e^{-\frac{T}{\alpha}}\right) - 2T\alpha e^{-\frac{T}{\alpha}} \cancel{+T^2 e^{-\frac{T}{\alpha}}}$$

$$Var(Z) = E[Z^2] - E[Z]^2 = \left(\cancel{2\alpha^2\left(1 - e^{-\frac{T}{\alpha}}\right)} + T^2 e^{-\frac{T}{\alpha}}\right) - \left(\alpha\left(1 - e^{-\frac{T}{\alpha}}\right) + T e^{-\frac{T}{\alpha}}\right)^2$$

$$2\left(\alpha^2 - \alpha^2 e^{-\frac{T}{\alpha}} - T\alpha e^{-\frac{T}{\alpha}}\right)$$

$$\Rightarrow Var(Z) = \alpha^2 - T\alpha e^{-\frac{T}{\alpha}} - \alpha^2 e^{-\frac{2T}{\alpha}} + \alpha^2 e^{-\frac{T}{\alpha}}$$